Instrumentation for Future Parallel Computing Systems

Edited by

Margaret Simmons
Rebecca Koskela
Ingrid Bucher

acm
PRESS

Frontier Series

Instrumentation for Future Parallel Computing Systems

Edited by

Margaret Simmons
Rebecca Koskela
Ingrid Bucher

ACM Press
New York, New Yo

Addison-Wesley Publishing Company
The Advanced Book Program
Redwood City, California • Menlo Park, California • Reading, Massachusetts
New York • Don Mills, Ontario • Wokingham, United Kingdom • Amsterdam
Bonn • Sydney • Singapore • Tokyo • Madrid • San Juan

ACM Press Frontier Series
Instrumentation for Future Parallel Computer Systems

First printed 1989

Library of Congress Cataloging-in-Publication Data

Instrumentation for future parallel computer systems / [edited by]
Margaret Simmons, Ingrid Bucher, Rebecca Koskela.
p. cm. — (ACM Press frontier series)
"Based on a Los Alamos National Laboratory workshop conducted in 1988"
—T.p. verso.
1. Parallel processing (Electronic computers) 2. Electronic digital
computers—Evaluation. 3. Computer software—Evaluation. I. Simmons,
Margaret, II. Bucher, Ingrid. III. Koskela, Rebecca. IV. Los Alamos National
Laboratory.
QA76.5.I4875 1989 004'.35—dc 19 89-302
ISBN 0-201-50390-5

ABCDEFGHIJ-AL-89

Instrumentation for Future Parallel Computing Systems

ACM PRESS

ACM Press books represent a collaboration between the Association for Computing Machinery (ACM) and Addison-Wesley Publishing Company to develop and publish a broad range of new works. These works generally fall into one of four series.

Frontier Series. Books focused on novel and exploratory material at the leading edge of computer science and practice.

Anthology Series. Collected works of general interest to computer professionals and/or society at large.

Tutorial Series. Introductory books to help nonspecialists quickly grasp either the general concepts or the needed details of some specific topic.

History Series. Books documenting past developments in the field and linking them to the present.

In addition, ACM Press books include selected conference and workshop proceedings.

CONTRIBUTORS

William Bohm
University of Manchester

William C. Brantley
IBM Thomas Watson Research Center

Robert J. Carpenter
National Bureau of Standards

Frederica Darema
IBM Thomas Watson Research Center

Blaine Gaither
Gould Electronics

Harry F. Jordan
University of Colorado

Koji Kinoshita
NEC Corporation

Allen D. Malony
Computer Science Department, University of Illinois

David Mandell
Los Alamos National Laboratory

Kevin P. McAuliffe
IBM Thomas Watson Research Center

James R. McGraw
Lawrence Livermore National Laboratory

Harry Nelson
Lawrence Livermore National Laboratory

Ton A. Ngo
IBM Thomas Watson Research Center

Gregory M. Papadopoulos
Massachusetts Institute of Technology

Daniel A. Reed
Computer Science Department, University of Illinois

Matthew Reilly
Digital Equipment Corporation

Patricia J. Teller
Department of Computer Science, The Courant Institute

Harold Trease
Los Alamos National Laboratory

PREFACE

In May, 1988 a "Workshop on Instrumentation for Future Parallel Computing Systems" was held in Santa Fe, New Mexico. The workshop was organized to exchange ideas among designers of hardware and software for high performance parallel computer systems, users of such systems, and members of the research community. Performance evaluation of parallel computer systems is a complex process; the main challenge in this area is the development of the required performance evaluation methods and tools to keep pace with the explosion of new system designs brought about by rapid technological advances. The next generation of high performance machines will likely all be parallel supercomputers. In order to effectively evaluate the performance of the current and future workloads on these systems, new kinds of measurement techniques and data will be necessary. Performance instrumentation, both hardware and software, on existing parallel systems is inadequate at best, or nonexistent.

The goal of the workshop was to clarify the desirable and feasible features of instrumentation for high performance computer systems, and to determine what should therefore be included in future system designs. The availability of integrated basic hardware and software for the collection of performance data would not only teach us how to program such systems more effectively, but would also lead to the design of more effective systems. To achieve this goal a group of designers, researchers, and users of such systems were invited to two intense days of presentations and discussion sessions.

For the discussions, participants were divided into three working groups: those interested in multiprocessors with common memory, a group concerned with parallel computers with distributed memory, and a group evaluating the special needs of dataflow computers. Each group was asked to consider the following questions during their discussions:

1. What quantities are necessary to characterize the performance of parallel computing systems?
2. If you could choose an ideal list of quantities that you would like to see measured, what would you choose and what priority would you give to each of these choices?
3. Which of the above quantities must be measured simultaneously?
4. Measurement tools usually distort the behavior of the system being measured. How much distortion is tolerable and how can it be minimized or corrected for?

Among the results of the workshop were two lists: a minimum set of measurement tools that are a necessity for future systems, and an additional set of measurement tools that would be highly desirable beyond this minimum set. The interaction between designers and users brought about the realization that many needs are common to both groups. The participants felt that this interaction also further deepened their knowledge of and appreciation for each other's positions and problems. The hardware designers better understood the limitations of software and the software designers felt that they gained insight into the problems presented by hardware. Users and researchers had an opportunity to make their needs and wishes known to both groups.

This volume contains both the papers presented at the general sessions of the workshop and summaries of the working group discussions. The book should be of interest to hardware and software designers of parallel computing systems, to users looking for a deeper understanding of system performance and striving to improve that performance, and to advanced students who wish to gain insight into the performance evaluation of parallel computer systems.

The papers included in the book were selected for their relevance to the questions posed by the workshop. The breadth of interest and expertise represented at the workshop is reflected in the range of papers chosen. The collection as a whole provides a description of the challenges and opportunities facing each of the groups of workshop attendees: designers, researchers, and users. The papers are organized into roughly two groups: those with proposed or existing instrumentation systems, and those relating experiences with such systems.

The first three chapters in the book present are keynote addresses that give details of the two best instrumented research systems in existence today: the Cedar project at the University of Illinois and the RP3 project at IBM. The first chapter relates the approaches taken in the implementation of instrumentation for the multiprocessor Cedar. One of the goals of the Cedar project was to build a prototype of a "performance instrumented computer" that offers the user the ability to observe parallel operation at various levels of detail. The next two chapters are concerned with the IBM multiprocessor project, RP3. Chapter 2 describes the performance monitoring hardware of the RP3. The RP3 system supports both a shared-memory and a message-passing paradigm. Performance monitoring instrumentation was included in the design phase of the machine in order to satisfy the primary goal of understanding parallel hardware and software. A methodology for analyzing the performance of parallel applications is presented in Chapter 3. Performance analysis of parallel applications is considered from two points of view: that of the application programmer and that of the hardware designer.

Chapter 4 discusses the visualization tools and design approach for performance instrumentation in the Cedar project. This chapter addresses the problems posed by the volume of data from performance measurements. These data must be presented in ways that emphasize important events and elide irrelevant data.

Chapter 5 presents an intriguing proposal to those who desire better instrumentation for debugging, program analysis, and performance tuning. Papadopoulos suggests that the perceived need for dynamic and specific instrumentation "is an artifact of the mismatch (and inadequacy) of current parallel programming languages and processor architecture." He proposes that determinate debugging, theoretical performance prediction, and relaxed execution instrumentation can be accomplished through declarative languages and dataflow machines.

In response to the first question posed to the working groups, Chapter 6 lists the minimum set of instrumentation that Gaither finds necessary for shared-memory parallel systems.

Bohm, et. al. state in Chapter 7 that the key problem inhibiting effective use of parallel computer systems is the absence of tools to support the process of mapping an application onto the parallel resources of a machine. They present the "levels of abstraction" view of parallel computer systems that helps to organize and structure an understanding of these systems and to categorize the debugging and monitoring tools needed to visualize program and machine behavior.

Chapter 8 describes the implementation of measurement support for software performance tuning on the 32 processor, shared-memory M31 VAX. It also discusses the problem of distortion posed in the fourth question to the working groups.

Carpenter presents a range of approaches to performance measurements in Chapter 9. This chapter also addresses the problem of the perturbations caused by performance measurements.

The next four chapters discuss performance instrumentation from the users' point of view. Chapter 10 discusses the instrumentation necessary for measuring barrier performance to determine the effect on the performance of a parallel program. Barriers are a convenient synchronization mechanism for parallel programs. In Chapter 11, Nelson describes his experiences with the hardware performance monitors on the CRAY X-MP at Lawrence Livermore National Laboratory. Chapter 12 discusses the problems encountered in parallelizing a large scientific code orginally written for a serial machine. Mandell and Trease list the software tools that would help in the development of large parallel codes. In Chapter 13 Kinoshita describes the ANALYZER/SX, a performance tuning tool available on the NEC series of supercomputers. This tool provides both a static and dynamic analysis of a computer program. He discusses program optimization through vectorization and tuning.

The common memory working group focused their discussions in two areas: measurement and tuning of applications' performance on current and future systems, and the performance analysis of hardware and software to identify bottlenecks in both the computer system and the application. Chapter 14 considers data available from existing instrumentation and the information that is currently not available in these two contexts.

In Chapter 15, Reed begins with a review of distributed-memory systems, including proposed communication paradigms and commercial systems. He then covers the issues relevant to distributed-memory performance instrumentation from the working group's perspective.

Dataflow architectures have major consequences for performance measurement and evaluation. Chapter 16 covers the discussions concerned with performance techniques and tools most appropriate to dataflow computers. McGraw gives some brief background information on dataflow computing and goes on to address the issues raised and to present possible techniques for performance evaluation of dataflow systems.

ACKNOWLEDGMENTS

Many people contributed to the success of the workshop and to the production of this book. We offer our sincere thanks to all of them.

In particular, we thank Ann Hayes, Andy White, and Norm Morse for their support and encouragement, Yvonne Martinez, Debbie Martinez and Sheila Girard, and especially Jan Hull for help with the organization of the workshop. The workshop was sponsored by Los Alamos National Laboratory. We would also like to thank Cray Research, Incorporated and Gould Electronics for their financial support.

A number of people have contributed to the preparation of this book, including Ann Garnett, Kyle Wheeler, Marge Blackwell, Lucille Maestas, and JoAnn Olivas. We are grateful for their help. We thank Pat Teller, Dan Reed, and Jim McGraw for their excellent summaries of the three working group discussions. They each spent many hours with audio tapes and transcripts trying to make sense and order out of confusion and chaos. They have done a superb job.

Finally, Yvonne Martinez has done an outstanding job in her careful and meticulous preparation of the manuscript. For this, we owe her special thanks.

FOREWORD

Time. Where *does* it go? We often ask ourselves this question, partly in the hopes of catching some of the elusive quantity that slips by us and partly to do better in utilizing it the next day. Those of us in the business of supercomputers can justifiably ask the same question and for the same reasons as we try to obtain optimal performance out of the machines and the applications that are executed on them. Where does the time go? Where is it spent?

In an attempt to get a better handle on this process of measuring supercomputer performance, we sponsored a workshop on instrumentation of supercomputers. Experts in hardware and software aspects, as well as major users of supercomputers, were invited to attend. Both the current status of instrumentation monitoring and needs for the future were discussed. This book represents the proceedings of that first workshop, both in the form of presented papers and synopses of workshop discussion groups.

February 1989 Ann H. Hayes

Instrumentation for Future Parallel Computing Systems

CONTENTS